Probability Theory Problems

Agnibho Roy

Summer 2020

I have taken the time to look at at some questions questions that deal with the generalized form of common probability problems and also some original ideas. I tried to make writing a question a warm-up exercise every day during the summer of 2020, and these are the manifestations of my creativity. Enjoy!

Contents

1 Coins

1.1 Alternating Sequences

1.2 Most Likely

2 Balls and Die

2.1 Colored Balls in Bags

2.2 Many Rolls of Die

2.3 Successive Numbers

Problem Statement

2.4 Closest Roll

Two people each bids a number before throwing a 30 faced die. Whoever gets closer to the number wins and wins the amount of money equal to the number they throw. e.g I bid 15 and you bid 16. the die lands on 10 then i win 10 from you. What's the best strategy and the expected payoff.

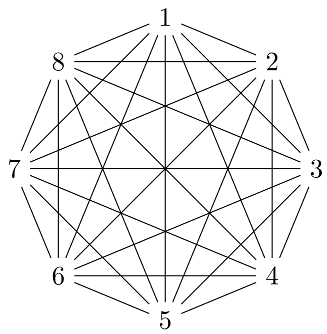
3 Walking Ants

All of the questions in this section have are related to variations on the ants taking a random walk on a surface type questions. I have found these sorts of questions very common in quantitative finance interviews, and some of these are inspired by a class that I took at Berkeley.

3.1 Returning on a Complete Graph

Problem Statement

We discuss an interesting problem where we consider an ant walking on a complete graph with n nodes. Let us say that the ant starts at some arbitrary node, and we are interested in finding the probability that after some number of steps, the ant returns to its starting node. More generally, the problem can be stated as follows: An ant travels on a complete graph and has an equal probability at every time step to travel to each its neighboring nodes. For a complete graph on n nodes, find the probability that, after k time steps, the ant returns to the node that it started on. Here is a visualization of the ant traveling on a complete graph of eight nodes starting at node 1:



Solution

The way to approach this problem would be to consider when it is possible for the ant to be at its original node, lets say some node v_1 WLOG, at the kth time step. Only if at the previous time step $k-1$, the ant is not at node v_1. If the ant was at any other node, then it would have a probability of $\frac{1}{n-1}$ of coming to the original node v_1. As a result, if the probability of returning at time step k is denoted as p_k, then it can be written recursively as:

$$p_k = \frac{1}{n-1}(1-p_{k-1}) = \frac{1}{n-1}(1-\frac{1}{n-1}(1-\dots p_0)$$

Where the initial condition is $p_0 = 1$. If we expand this further, we see that we see a very structured sequence, which we can write as:

$$\begin{aligned}
p_k &= \frac{1}{n-1} - (\frac{1}{n-1})^2 + \cdots + (\frac{1}{n-1})^k - (\frac{1}{n-1})^k \\
&= \frac{1}{n-1} - (\frac{1}{n-1})^2 + \cdots + (\frac{1}{n-1})^{k-1} \\
&= \sum_{i=1}^{k-1} -(-1)^k \frac{1}{n-1}^i \\
&= \frac{\frac{1}{n-1}((-\frac{1}{n-1})^{k-1} - 1)}{-\frac{1}{n-1} - 1} \\
&= -\frac{1}{n}((-\frac{1}{n-1})^{k-1} - 1)
\end{aligned}$$

Where the last equality is using the sum of an infinite series with common ratio $-\frac{1}{n-1}$ and initial value $\frac{1}{n-1}$.

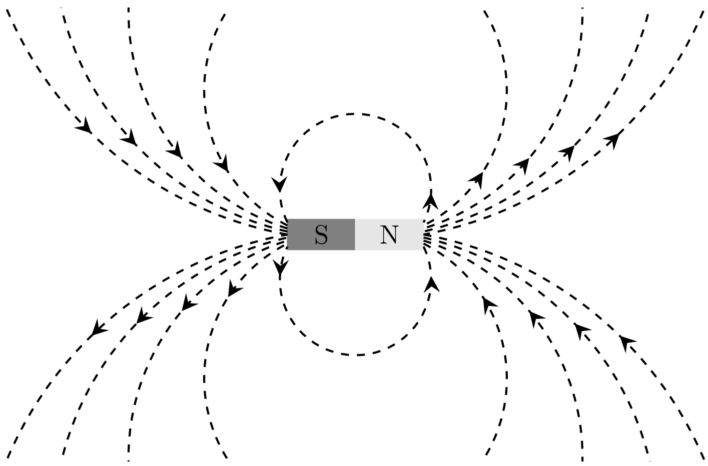
Key Takeaways

Recursion is very powerful in probability. When we are asked a question that is generalized to n steps, we should express that question in terms of the previous step $n-1$, and pretend that the problem is already solved. After unraveling the sequence, most times we can get to a closed-form expression.

3.2 Multiple Paths

Problem Statement

Let us say that a magnetic ant is thrown into a magnetic field and as a result, is randomly thrown across the field along one of the curves (pictured below). As one can imagine, there are an infinite amount of field curves, but it is less likely that the ant gets thrown into a field curve that is further away from the magnet; however, if the ant does, then it takes a longer time to return back.



Specifically, denote a_i as the ith closest arc to the magnet from the top, and b_i as the ith closest arc to the magnet from the bottom. The probability that the ant is thrown into arc a_i or arc b_i is $\frac{1}{4^i}$, and the amount of time that it will take to return is 2^i. The remaining probability goes towards being able to escape the magnetic field through the north or south poles, which happens with probability $\frac{1}{6}$ each (you can convince yourself that the sum of all of these probabilities is 1). The amount of time it takes to escape is 1 hour. Find the expected time that the magnetic ant will spend in the magnetic field.

Solution

The way to approach this problem is to first consider the what we are looking for: the expected total time that the magnetic ant will spend in the magnetic field, and write that expectation as a summation of cases, where the cases are getting thrown into each of the n arcs as $n \to \infty$, weighed by the probability of getting released into each of the arcs.

Think about it like this: what if get thrown into the first arc? First, we enter this arc with probability $\frac{1}{4^1} = \frac{1}{4}$ and we will take $2^1 = 2$ time to get back, but after we come back, we are right back where we started. This is just one of the cases, but this can be repeated in the same manner for the rest of the arcs. For when we escape, which happens with probability $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$, we just take an hour to escape and we are done. In total, the expectation can be written recursively as something like this (where β is the expectation):

$$\beta = \frac{1}{4}(2+\beta) + \ldots (\text{other cases}) \cdots + \frac{1}{3}(1)$$

We seem to have a nice sequence for the probabilities and return times for each of the arcs, so we can write it as a summation and solve for β:

$$\begin{aligned}
\beta &= \sum_{i=1}^{\infty}(\frac{2^i}{4^i} + \frac{\beta}{4^i}) + \frac{1}{3} \\
&= \sum_{i=1}^{\infty}(\frac{1}{2})^i + \beta\sum_{i=1}^{\infty}(\frac{1}{4})^i + \frac{1}{3} \\
&= 1 + \frac{\beta}{3} + \frac{1}{3} \\
\frac{2}{3}\beta &= \frac{4}{3} \\
\beta &= 2
\end{aligned}$$

Where the summations were simplified by using properties of an infinite geometric series. As a result, the magnetic ant is expected to spend 2 hours in the magnetic field before escaping. As an additional exercise, I would suggest exploring different values for the probabilities of traveling along the arcs to see when the expected value converges or diverges (ex. think about if the probabilities of going along each arc is instead $\frac{1}{2^i}$ instead of $\frac{1}{4^i}$)

Key Takeaways

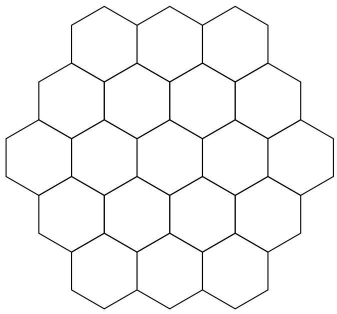
Think about all the different cases of a problem and see if you can express them in terms of your current position.

3.3 Escaping a Honeycomb

Source: Jane Street Bonus Problem 3 (More on Real Number Bonus Problems)

Problem Statement

Imagine an ant is stuck inside a beehive that is hexagonal, consisting of cells that are 19 smaller hexagons (depicted below). Imagine that the ant starts in the middle cell and since he does not know what cell within the beehive he is in at any moment in time, he just uniformly chooses one of the six cells around him. Find the expected time for the ant to leave the beehive. Imagine that when the ant reaches the end, there are imaginary cells in place of empty spaces outside the beehive to find the probability that the ant will leave.



Solution

This question almost begs to be structured into a markov chain, but we can simplify the process by taking advantage of the symmetry of the beehive structure. Symmetry is just another way to collapse multiple states into one if they behave in the same way. For example, in the first "ring" of cells outside of the middle, we see that they all have an equal probability of leaving to the outer ring, staying in the same ring, or moving into the middle. When we see this, we can collapse the six cells in the ring into a singular state in order to reduce the number of states in our markov chain, and repeat this for all of the rings What this implies in terms of the question, if we are in a cell, then we have the same expectation of leaving as any other cell in the same ring. The markov chain would be:

Key Takeaways

3.4 Speedy Skippers

Imagine an ant traveling in one dimension from city to city, where there are a total of N cities, and takes the ant train to do so. He needs to deliver a package on the way to the last city in the path, and this package is at city k. He needs to be fast however, and so this train takes him to each of the cities ahead of him with uniform probability (in other words, if the ant is at city l, then the train will go next to any of the cities ahead of him each with probability $\frac{1}{N-l}$). What is the probability that the train will stop at city k to allow him to pick up his package on his journey?